

116TH CONGRESS
2D SESSION

S. 4990

To require the Office for Civil Rights of the Department of Health and Human Services to conduct a study and issue a report on the de-identification of data pursuant to privacy regulations.

IN THE SENATE OF THE UNITED STATES

DECEMBER 9, 2020

Mr. CASSIDY introduced the following bill; which was read twice and referred to the Committee on Health, Education, Labor, and Pensions

A BILL

To require the Office for Civil Rights of the Department of Health and Human Services to conduct a study and issue a report on the de-identification of data pursuant to privacy regulations.

1 *Be it enacted by the Senate and House of Representa-*

2 *tives of the United States of America in Congress assembled,*

3 **SECTION 1. SHORT TITLE.**

4 This Act may be cited as the “Health Data De-identi-

5 fication Report Act”.

1 **SEC. 2. STUDY AND REPORT BY THE HHS OFFICE FOR CIVIL
2 RIGHTS.**

3 (a) IN GENERAL.—The Office for Civil Rights of the
4 Department of Health and Human Services, in coopera-
5 tion with the National Institute of Standards and Tech-
6 nology, shall conduct a study that considers and compares
7 the effectiveness and validity of safe harbor and expert
8 determination as methods of carrying out the de-identi-
9 fication of data required under section 164.514 of title 45,
10 Code of Federal Regulations (as in effect on the date of
11 enactment of this Act), and shall submit to Congress a
12 report on such study not later than 270 days of the date
13 of enactment of this Act.

14 (b) CONTENT OF STUDY AND REPORT.—The study
15 and report under subsection (a) shall consider—

16 (1) any known instances where data that was
17 de-identified as required under section 164.514 of
18 title 45, Code of Federal Regulations (as in effect on
19 the date of enactment of this Act) became re-identi-
20 fied, even partially, and, with respect to each such
21 instance, how such re-identification occurred and
22 any harm incurred by any individual whose data was
23 re-identified;

24 (2) the frequency by which each of the safe har-
25 bor method and the expert determination method is
26 used for the de-identification of data as described in

1 subsection (a), and the sizes of the data sets on
2 which each such method is used;

3 (3) the costs of each such method of de-identi-
4 fication, and the relative utility of the data that is
5 de-identified through each such method;

6 (4) which of the 2 methods of de-identification
7 renders data less likely to be easily re-identifiable;

8 (5) whether either such method removes infor-
9 mation that would help address health care dispa-
10 rities, thereby making it more difficult to address such
11 disparities;

12 (6) which of the 2 methods is most commonly
13 used for purposes of medical research, and the bene-
14 fits and disadvantages of using each such method
15 for such purpose;

16 (7) the risk of re-identification by each of the
17 2 methods, especially when the resulting de-identi-
18 fied dataset will be accessible by entities that have
19 access to additional data;

20 (8) whether or not there are use cases that can-
21 not be performed using data de-identified by either
22 of the 2 methods; and

1 (9) the ease or difficulty of expert determina-
2 tion, taking into consideration cost, accessibility, and
3 qualifications of experts.

